

ANUARIO DEL SEMINARIO DE FILOLOGÍA
VASCA «JULIO DE URQUIJO»

International Journal of Basque Linguistics and Philology

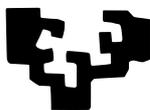
LII: 1-2 (2018)

*Studia Philologica et Diachronica
in honorem Joakin Gorrotxategi*

Vasconica et Aquitanica

Joseba A. Lakarra - Blanca Urgell
(arg. / eds.)

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

How Many Language Families are there in the World?

Lyle Campbell

University of Hawai'i Mānoa

Abstract

The question of how many language families there are in the world is addressed here. The reasons for why it has been so difficult to answer this question are explored. The answer arrived at here is 406 independent language families (including language isolates); however, this number is relative, and factors that prevent us from arriving at a definitive number for the world's language families are discussed. A full list of the generally accepted language families is presented, which eliminates from consideration unclassified (unclassifiable) languages, pidgin and creole languages, sign languages, languages of undeciphered writing systems, among other things. A number of theoretical and methodological issues fundamental to historical linguistics are discussed that have impacted interpretations both of how language families are established and of particular languages families, both of which have implications for the ultimate number of language families.

Keywords: *language family, family tree, unclassified language, uncontacted groups, language isolate, undeciphered script, language surrogate.*

1. Introduction

How many language families are there in the world? Surprisingly, most linguists do not know. Estimates range from one (according to supporters of Proto-World) to as many as about 500. The goals of this paper are to address the question of how many language families there are in the world, and to consider what makes this question so difficult to answer. To break any suspense and to anticipate the conclusion here at the outset, let me report the answer here: 406 independent language families (including language isolates). However, this number is far from straightforward or conclusive, as we shall see.

2. What does not count as a language family?

It will be helpful to begin by eliminating things that do not count for determining the number of language families. By convention, the following are not included in the lists of the world's language families.

2.1. Language surrogates

Excluded as language surrogates are whistled, drum(med) languages, codes (that convert language information into some other form of representation, e.g. Morse Code), and sign languages. In the literature on sign languages, "families" of sign languages are treated, but these are not counted in the overall number of language families in the world. This may seem unfortunate, but it is not inappropriate since the term "family" for sign languages typically describes a contact relationship rather than shared ancestry. Relatedness among sign languages is not determined, as with spoken languages, by showing descent from a common ancestor through the application of the comparative method. Rather, a sign language's membership in a "family" of sign languages reflects the sign language which has most influenced it; the "family" relationship is more about language contact (lateral transmission) and is not about descent from a common ancestor. For example, American Sign Language (ASL) is considered to belong to the French Sign Language (LSF) family, but ASL is not really a descendant of Old French Sign Language, but rather is the result of mixing of indigenous signs in the first American deaf schools with foreign French signs brought by teachers (Clark 2017: 13).

2.2. Pidgin and creole languages

Traditionally, it has been held that pidgins are minimal contact languages used among groups for whom it is no one's native language, and that creoles develop from pidgins when they acquired native speakers (when children of parents who use pidgin grow up with pidgin as their first language). Creoles are said to stem from multiple parent languages, with most of the lexicon from a lexifier language and parts of the grammar from a 'substrate' languages. In this view, creoles present a problem for classification, since genealogical classification of languages permits one and only one parent per language. Thomason and Kaufman's (1988) proposed solution to the problem was that only languages that have a single parent qualify for classification into language families. Creoles, in this view, are not changed forms of a single parent language; they do not arise in direct transmission from one speaker to another and they descend from more than one language, and are therefore disqualified from classification in terms of language families.

This traditional view of the origin of pidgins and creoles has been challenged in recent times by a number of scholars. Some do not believe that a creole must evolve from a prior pidgin. In their view, pidgins came about in trade contact among people who keep their native languages for everyday communication, while creoles developed in a different way, in colonies from non-standard varieties of mostly European languages whose speakers interacted intensely with plantation laborers and

slaves. The language they used remained essentially a version of the European language underlying it. In this view, each creole has a single ancestor. It is the language of the founder population, speakers of the dominant European language, that predominates in the formation of the creole language and in its content, and therefore creoles present no particular problem for genetic classification. In this view they are classified in the same way as any other language would be. So, for example, Jamaican Creole and Tok Pisin are Germanic languages, closely related to English, and Haitian Creole and Mauritian Creole are Romance, connected with French (see, for example, Mufwene 1996). No doubt controversy concerning the origins and classification of creole languages will continue. In any event, creoles either do not count because they have multiple ancestors, or they do count but present no significant complication because they are regular members of the language family to which their chief lexifier language belongs.

2.3. Unclassified languages

An unclassified language is one for which there are not enough data available to be able classify it – these languages lack enough data for them to be compared meaningfully with other languages and therefore their possible kinship remains unknown, though some are sometimes inaccurately listed with language isolates. For language isolates, sufficient data do exist to compare them with other languages; they are not grouped in larger genetic classifications with any other language because these comparisons do not support any linguistic kinship with any other language.

There are a good number of unclassified, indeed unclassifiable, languages whose genetic affiliation is unknown, of two sorts.

The first are the extinct languages that are too poorly attested to be grouped with any other language or language family. A few examples include:

- Aranama-Tamique, Texas
- Baenan, Brazil
- Camunico, Northeast Italy (survived to 2nd half of 1st millennium BCE)
- Eteocretan, Crete, 7-3 centuries BCE (cf. Michalowski 2018)
- Gamela, Brazil
- Gule, Sudan
- Kaskean, Northeast Anatolia 2nd millennium BCE
- Maratino, Mexico
- Minkin, Australia
- Mure, Bolivia
- Naolan, Mexico
- Pictish, Scotland 7-10 centuries CE, few inscriptions
- Solano, Texas, Mexico
- Sorothaptic, Iberian Peninsula, pre-Celtic, Bronze Age (cf. Michalowski 2018)
- Tarairiú, Brazil

(For several other unclassified extinct languages of Asia and Europe, see Michalowski 2018; for others of South America, see Campbell 2012.)

The second kind of unclassified languages consists of extant, still-spoken languages that cannot be classified for lack of data, i.e., languages not yet described suf-

ficiently to be able to compare them meaningfully with other languages in order to determine whether they may have relatives. A few examples are:

In Africa: Bung, Lufu, Kujargé, and perhaps Mpre (Mpra) (see Blench 2018).

In Asia and the Pacific: Sentinelese, Bhatola, Waxianghua, Turumsa, Kembra, and Lepki.

In South America there are many, for example: Ewarhuyana, Kaimbé, Kambiwá, Kapinawá, Pankararé, Truká, Tremembé, Wakoná, Wasu, etc. (See Campbell 2012; cf. Seifart and Hammarström 2018).

It should be noted that some of these unclassified languages could also turn out to be language isolates and thus increase the world's total number of language families, if we had more data on them. However, without data we cannot know.

2.4. “Languages” known only from undeciphered scripts

For undeciphered scripts often there is speculation about what the language represented in a particular script may have been. However, usually we do not know what language the undeciphered scripts may represent. Thus the languages of these scripts are also “unclassified.” They could represent known languages, or previously unknown languages that nevertheless belong to known language families, or they could belong to as yet unknown language families. Until and unless these scripts are deciphered, we will not know whether additional, previously uncounted language families are involved.

Some examples of undeciphered scripts are: the Byblos Syllabary, Cretan Hieroglyphs, Cypro-Minoan, Linear A (Minoan), Linear Elamite, and Proto-Elamite (tablets). (See Michalowski 2018 for discussion of several of these.)

2.5. Constructed languages (artificial languages)

Also not included in lists of language families are constructed languages such as Esperanto, Ido, Interlingua, Volapük; Quenya, Sindarin; Klingon, Na'vi, Dothraki, Parseltongue, etc.

2.6. Glossolalia (“speaking in tongues”)

In glossolalia people appear to speak in languages unknown to them, involving incomprehensible “speech” that is often associated with trance states and religious ecstasy. Although glossolalia is often associated with specific varieties of Christianity, speaking in tongues is attested in various other religions around the world. These are not real languages by linguistic criteria.

2.7. Animal communication systems

All species communicate, in some fashion or another, visual, auditory, or olfactory, but only human language counts for language families.

3. What does count as a language family?

A language family is the set of languages for which there is sufficient evidence to show that they descend from a common ancestor. Language isolates are also language families, but have only one member language, not related to any other language or language family. The ultimate number of language families in the world is the set of independent families (including isolates) that cannot reliably be shown to be related to any other language family.

The following represents what approaches a consensus among experts of the language families in the historical linguistics of the different regions.

Africa (42)

- | | | |
|---------------------------------|-----------------------|--|
| 1. Afro-Asiatic | 16. Khoe | 33. Saharan |
| 2. Bangi Me (isolate) | 17. Kresh-Aja | 34. Sandawe (isolate?) |
| 3. Berta | 18. Kx'a (Ju + †Huan) | 35. Songhay |
| 4. Central Sudanic | 19. Koman | 36. South Omotic
(Aroid?) |
| 5. Daju | 20. Kuliak | 37. Surmic |
| 6. Dizoid | 21. Kunama | 38. Ta-Ne-Omotic |
| 7. Dogon | 22. Laal (isolate) | 39. Tama (Taman) |
| 8. Eastern Jebel | 23. Maban | 40. Tegem (Lafofa)
(isolate?,
unclassified?, family?,
Niger-Congo?) |
| 9. Furan | 24. Mande | 41. Temein |
| 10. Gimojan (Gonga-
Gimojan) | 25. Mao | 42. Tuu |
| 11. Hadza (isolate) | 26. Nara (isolate?) | |
| 12. Heiban | 27. Narrow Talodi | |
| 13. Ijoid | 28. Niger-Congo | |
| 14. Jala*(?) (isolate) | 29. Nilotic | |
| 15. Kadu (Kadugli-
Krongo) | 30. Nubian(+Meroitic) | |
| | 31. Nyimang | |
| | 32. Rashad | |

North America (54)

- | | | |
|--------------------------------------|------------------------------------|---|
| 1. Adai* (isolate [unclassified?]) | 12. Chumashan* | 24. Karuk (Karok) (isolate) |
| 2. Algic | 13. Coahuilteco* (isolate) | 25. Keresan |
| 3. Alsea* (isolate) | 14. Cochimí-Yuman | 26. Kiowa-Tanoan |
| 4. Atakapa* (isolate, small family?) | 15. Comecrudan* | 27. Kootenai (Kutenai) (isolate) |
| 5. Beothuk* (isolate) | 16. Coosan* | 28. Maiduan |
| 6. Caddoan | 17. Cotoname* (isolate) | 29. Muskogean |
| 7. Cayuse* (isolate) | 18. Eskimo-Aleut | 30. Na-dene (strict sense, Athapaskan-Eyak—Tlingit) |
| 8. Chimakuan* | 19. Esselen* (isolate) | 31. Natchez* (isolate) |
| 9. Chimariko* (isolate) | 20. Haida (isolate, small family?) | 32. Palaihnihan |
| 10. Chinookan* | 21. Iroquoian | |
| 11. Chitimacha* (isolate) | 22. Kalapuyan* | |
| | 23. Karankawa* (isolate) | |

- | | | |
|--------------------------------|-----------------------------|---------------------|
| 33. Plateau (Plateau Penutian) | 40. Takelma* (isolate) | 47. Wakashan |
| 34. Pomoan | 41. Timucuan* | 48. Washo (isolate) |
| 35. Salinan* | 42. Tonkawa* (isolate) | 49. Wintuan |
| 36. Salishan | 43. Tsimshianic | 50. Yana* (isolate) |
| 37. Shastan* | 44. Tunica* (isolate) | 51. Yokutsan |
| 38. Siouan-Catawban | 45. Utian (Miwok-Costanoan) | 52. Yuchi (isolate) |
| 39. Siuslaw* (isolate) | 46. Uto-Aztecan | 53. Yukian* |
| | | 54. Zuni (isolate) |

Mexico and Mesoamerica (14)

- | | | |
|-------------------------|------------------------------------|--------------------|
| 1. Cuitlatec* (isolate) | 7. Misumalpan | 12. Tequistlatecan |
| 2. Guaicurian* | 8. Mixe-Zoquean | 13. Totonacan |
| 3. Huave (isolate) | 9. Otomanguean | 14. Xinkan(*) |
| 4. Jicaquean (Tol) | 10. Seri (isolate) | |
| 5. Lencan* | 11. Tarascan (Purépecha) (isolate) | |
| 6. Mayan | | |

South America (107)

- | | | |
|--|---|------------------------------------|
| 1. Aikaná (isolate) | 17. Candoshi (Canndoshi-Sharpa) (isolate) | 34. Guachi* (isolate) |
| 2. Andaquí* (isolate) | | 35. Guaicuruan |
| 3. Andoque (isolate) | | 36. Guajiboan |
| 4. Arara do Rio Branco* (Arara do Beiradão, Mato Grosso Arara) (isolate) | 18. Canichana* (isolate) | 37. Guamo* (isolate) |
| 5. Arawakan | 19. Cariban | 38. Guató* (isolate) |
| 6. Arawan | 20. Cayuvava* (Cayubaba) (isolate) | 39. Harákmbut-Katukinan |
| 7. Atacameño (Cunza, Kunza)* (isolate) | 21. Chapacuran | 40. Huarpean* |
| 8. Awaké (Arutani)* (isolate) | 22. Charruan* | 41. Irantxe (Müncü) (isolate) |
| 9. Aymaran | 23. Chibchan | 42. Itonama (isolate) |
| 10. Barbacoan | 24. Chipaya-Uru | 43. Jabutian |
| 11. Betoí-Jirara* (isolate) | 25. Chiquitano (isolate) | 44. Jêan (Jê family) |
| 12. Boran | 26. Chocoan | 45. Jeikó* (isolate) [Macro-Jêan?] |
| 13. Bororoan | 27. Cholonan* | 46. Jirajaran* |
| 14. Cahuapanan | 28. Chonan | 47. Jivaroan |
| 15. Camsá (isolate) | 29. Chono* (isolate) | 48. Jotí (Yuwana) (isolate) |
| 16. C a ñ a r - P u r u h á (Ecuador) (uncertain family of 2 languages) | 30. Cofán (A'ingá) (isolate) | 49. Kakua-Nukak |
| | 31. Culli (Culle)* (isolate) | 50. Kamakanan* |
| | 32. Esmeralda* (Atacame) (isolate) | 51. Kapixaná (Kanoé) (isolate) |
| | 33. Fulnio (Yaté) (isolate) | 52. Karajá |

- | | | |
|--|---------------------------------------|-----------------------------------|
| 53. Karirían* | 70. Nadehup
("Makúan") | 87. Taushiro (isolate) |
| 54. Kaweskaran | 71. Nambiquaran | 88. Tequiraca* (isolate) |
| 55. Krenákan
(Botocudan) | 72. Ofayé (Opayé)
(isolate) | 89. Tikuna-Yurí |
| 56. Kwaza (Koayá)
(isolate) | 73. Omurano*
(isolate) | 90. Timotean* |
| 57. Leco* (isolate) | 74. Otomacoan* | 91. Tiniguan* |
| 58. Lule-Vilelan* | 75. Paez (isolate?) | 92. Trumai (isolate) |
| 59. Máko* (Maku)
(isolate) | 76. Pano-Takanan | 93. Tukanooan |
| 60. Mapudungun | 77. Payaguá* (isolate) | 94. Tupían |
| 61. Mascoyan (Enlhet-
Enenlhet) | 78. Puinave (isolate) | 95. Urarina (isolate) |
| 62. Matacoan | 79. Puquina* (isolate) | 96. Waorani (isolate) |
| 63. Matanawí* (isolate) | 80. Purí-Coroado*
(isolate) | 97. Warao (isolate) |
| 64. Maxakalían | 81. Quechuan | 98. Witotoan |
| 65. Mochica (Yunga)*
(isolate) | 82. Rikbaktsá
(Canoeiro) (isolate) | 99. Xukurúan* |
| 66. Mosestén-Chinamé
(isolate) | 83. Sáliban | 100. Yagan (Yámana)*
(isolate) |
| 67. Movima (isolate) | 84. Sapé (Kaliana)*
(isolate) | 101. Yaguan |
| 68. Munichí* (isolate) | 85. Sechura-Catacaoan* | 102. Yanomaman |
| 69. Muran (Pirahã)
(isolate, small family?) | 86. Taruma* (Taruamá)
(isolate) | 103. Yaruro (Pumé)
(isolate) |
| | | 104. Yuracaré (isolate) |
| | | 105. Yurumangui*
(isolate) |
| | | 106. Zamucoan |
| | | 107. Zaparoan |

Eurasia (30)

- | | | |
|------------------------------------|---|---------------------------------------|
| 1. Ainu* (isolate) | 11. Japonic | 21. Northwest Caucasian |
| 2. Basque (isolate) | 12. Kartvelian | 22. Sino-Tibetan |
| 3. Burushaski (isolate) | 13. Kassite* (isolate) | 23. Sumerian* (isolate) |
| 4. Chukotko-
Kamchatkan | 14. Koreanic | 24. Tai-Kadai |
| 5. Dravidian | 15. Kusunda (isolate) | 25. Tungusic |
| 6. Elamite* (isolate) | 16. Miao-Yao (Hmong-
Mien) | 26. Turkic |
| 7. Hattic* (isolate) | 17. Mongolian | 27. Tyrsenian (Etruscan-
Lemnian)* |
| 8. Hruso (Hruso-Aka)
(isolate?) | 18. Nakh-Dagestanian
(Northeast Caucasian) | 28. Uralic |
| 9. Hurrian-(Hurro-
Uratian)* | 19. Nihali (isolate) | 29. Yeniseian |
| 10. Indo-European | 20. Nivkh (isolate, possi-
bly a small family) | 30. Yukaghir |

Pacific (129)

- | | | |
|--------------------------|--|------------------------------|
| 1. Abinomn (isolate) | 5. Anêm (isolate) | bot, Kambrambo)
(isolate) |
| 2. Abun (isolate) | 6. Angan | |
| 3. Afra (Usku) (isolate) | 7. Anim | 9. Arafundi |
| 4. Amtó-Musan | 8. Ap Ma (Botin, Kam-
bot, Kambrambo) | 10. Asaba (isolate) |

- | | | |
|---------------------------------------|-----------------------------------|--------------------------------------|
| 11. Austroasiatic | 51. Kehu (isolate) | 92. Powle-Ma (“Molof”) (isolate) |
| 12. Austronesian | 52. Kibiri-Porome (isolate) | 93. Purari (“Namau”) (isolate) |
| 13. Awin-Pa | 53. Kimki (isolate) | 94. Pyu (isolate) |
| 14. Baibai-Fas | 54. Kiwaiian | 95. Sause (isolate) |
| 15. Baining | 55. Koiarian | 96. Savosavo (isolate) |
| 16. Baiyamo (isolate) | 56. Kol (isolate) | 97. Senagi |
| 17. Banaro (isolate) | 57. Kolopom | 98. Sentani |
| 18. Bayono-Awbono | 58. Konda-Yahadian | 99. Sepik |
| 19. Bilua (isolate) | 59. Kosare (isolate) | 100. Sko (Skou) |
| 20. Bogaya (isolate) | 60. Kuot (isolate) | 101. Somahai |
| 21. Border | 61. Kwalean | 102. South Bird’s Head |
| 22. Bosavi | 62. Kwerbic | 103. South Bougainville |
| 23. Bulaka River | 63. Kwomtari | 104. Suki-Gogodala |
| 24. Burmeso (isolate) | 64. Lakes Plain | 105. Sulka (isolate) |
| 25. Busa (Odiai) (isolate) | 65. Lavukaleve (isolate) | 106. Tabo (Waia) (isolate) |
| 26. Dagan | 66. Left May (Arai) | 107. Taiap (isolate) |
| 27. Damal (Uhunduni, Amung) (isolate) | 67. Lepki-Murkim | 108. Tambora* (isolate) |
| 28. Dem (isolate) | 68. Lower Sepik-Ramu | 109. Tanahmerah (isolate) |
| 29. Dibiyaso (isolate) | 69. Mailuan | 110. Taulil-Butam |
| 30. Doso-Turumsa | 70. Mairasi | 111. Teberan |
| 31. Duna (isolate) | 71. Manubaran | 112. Timor-Alor-Pantar |
| 32. East Bird’s Head | 72. Marori (Moraori) | 113. Tor-Orya |
| 33. East Kutubu | 73. Masep (isolate) | 114. Torricelli |
| 34. East Strickland | 74. Mawes (isolate) | 115. Touo (isolate) |
| 35. Eastern Trans-Fly | 75. Maybrat (isolate) | 116. Trans New Guinea |
| 36. Eleman | 76. Mombum (family, 2 languages) | 117. Turama-Kikori |
| 37. Elseng (Morwap) (isolate) | 77. Monumbo (family, 2 languages) | 118. Ulmapo (“Mongol-Langam”) |
| 38. Fasu (isolate) | 78. Mor (isolate) | 119. Walio |
| 39. Geelvink Bay | 79. Morehead-Wasur | 120. West Bird’s Head |
| 40. Goilalan | 80. Mpur (isolate) | 121. West Bomberai |
| 41. Great Andamanese | 81. Namla-Tofanma | 122. Wiru (isolate) |
| 42. Guriaso (isolate) | 82. Nimboran | 123. Yale (Yalë, Nagatman) (isolate) |
| 43. Hatam-Mansim | 83. North Bougainville | 124. Yareban |
| 44. Inanwatan | 84. North Halmahera | 125. Yawa |
| 45. Kaki Ae (isolate) | 85. Ndu | 126. Yele (Yéli Dnye) (isolate) |
| 46. Kamula (isolate) | 86. Onge-Jarawa | 127. Yerakai (isolate) |
| 47. Kapauri (isolate) (Kapori) | 87. Pahoturi | 128. Yetfa-Biksi (isolate) |
| 48. Karami | 88. Pauwasi | 129. Yuat |
| 49. Kaure-Narau (possibly an isolate) | 89. Pawaia | |
| 50. Kayagar | 90. Pele-Ata | |
| | 91. Piawi | |

Australia (30)

- | | | |
|--|-----------------------------|---|
| 1. Bachamal* (isolate, possibly North Daly family) | 12. Mangarrayi* (isolate) | 23. Southern Daly |
| 2. Bunaban | 13. Maningrida | 24. Tangkic |
| 3. Eastern Daly* | 14. Maran | 25. Tiwi (isolate) |
| 4. Gaagudju* (isolate) | 15. Marrku-Wurrugu | 26. Umbugarla/Ngurm-bur* (isolate or small family?) |
| 5. Garrwan | 16. Mirndi (Mindi) | 27. Wagiman (Wageman)* (isolate) |
| 6. Giimbiyu* | 17. Northeastern Tasmanian* | 28. Wardaman* (isolate or small family) |
| 7. Gunwinyguan | 18. Northern Daly | 29. Western Daly |
| 8. Iwaidjan | 19. Nyulnyulan | 30. Worrorran |
| 9. Jarrakan | 20. Oyster Bay* | |
| 10. Kungarakany* (isolate) | 21. Pama-Nyungan | |
| 11. Limilngan* | 22. Southeastern Tasmanian* | |

Thus, the total of language families in the world is c. 406.

Of these 406 language families, 94 are extinct – that is no language that is a member of these extinct families has any remaining native speakers. These extinct families are marked with an asterisk (*) after them in the lists above. This means that 23.2% of the linguistic diversity of the world, calculated in terms of language families, is now lost. Many other language families will also soon become extinct, as the last surviving languages in these families lose their last speakers.

However, this list is far from definitive, though close to representing consensus views for the various regions. Opinion varies even among specialists at times. One area of difference is that some scholars reject certain proposals of relatedness among certain languages that others find more convincing. (For discussion of the methods of establishing distant genetic relationships, see Campbell and Poser 2008.) Sometimes scholars differ over whether a language should be considered unclassified or whether there is enough attestation to classify it. When the data suggest that it cannot be related to any other language, then they consider it language isolate, which then increases the ultimate number of language families they would count. In what follows, for illustration's sake, I mention some of the cases that have been classified differently at times in the literature.

One example is *Aranama* (connected with Tamique), which shows up often in lists of the language families of North America. It was a language at the Franciscan mission of Espiritu Santo de Zúñiga, founded on the lower Guadalupe River in Texas in 1726. The entire attestation is one single word and one two-word phrase: *himiáyana* 'water' and *Himiána tsýi!* 'Give me water!', recorded by Albert Gatschet in 1884 from a Tonkawa man, Old Simon, who also gave a short vocabulary of Karankawa (another Texas language). Old Simon called this language *Hanáma* or *Háname* (Gatschet 1884); the only people from the area who had a name similar to this were known as the *Aranama*, *Saranames*, or *Jaranames*. So, with good reason, *Aranama* is now considered unclassified and is not listed among the language families of North America. (See Mithun 2018 for details.)

Calusa (of Florida) is also often listed as a language isolate, although, as Mithun (2018) explains, the language is known from only about a dozen words from 1575 from a Spanish captive among the Calusa and from 50-60 place names. Early accounts report that Calusa was distinct from other languages of the area. As Mithun (2018: 193) says, “with such a small record, however, its [Calusa’s] status cannot be confirmed.” It is best considered unclassified.

Cañar-Puruhá (Ecuador) is often listed as a language family with two members, Cañar (Cañari) and Puruhá (Puruguay). However, both languages are extinct with very little attestation on which to base comparisons (Adelaar and Muysken 2004: 396-7). The status of this putative family, therefore, is unclear.

Gumuz (in the Ethiopian-Sudan border area) is often treated as a branch of putative Nilo-Saharan, though this is unsubstantiated. It has also been considered a language isolate, though Blench (2018: 177) now considers it to be related to Koman languages (part of “Nilo-Saharan”).

Hruso (Hruso-Aka) of Arunachal Pradesh, India was long assumed to be a Tibeto-Burman language (Sino-Tibetan family), though it is now thought actually to be a language isolate. Words that Hruso (Hruso-Aka) shares with Tibeto-Burman languages appear to be loans, and there is insufficient evidence in the rest of the lexicon to support a Tibeto-Burman (Sino-Tibetan) affiliation. (cf. Glottolog, <http://glottolog.org/resource/languoid/id/hrus1242>, accessed 12-3-2017.)

The status of extinct *Kwadi* (of Angola) is uncertain. Güldemann (2004, 2008) argued that Kwadi is part of Khoe, i.e. Central Khoisan, although evidence for this is not straightforward. Blench (2018: 181) believes Kwadi’s affiliation cannot be resolved.

Kujarge (on the Chad-Sudan border), often listed as a language isolate, is best considered unclassified. Blench (2018: 181) mentions that “the exiguous nature of the dataset is to be regretted.” Since it is unlikely that more data will become available, Kujarge’s classification remains unresolved.

Long extinct *Meroitic* (from the First Cataract of the Nile to the Khartoum area in Sudan) was usually unclassified or often treated as a language isolate, though a number of hypotheses of proposed relationships were made, usually grouping it somehow with “Nilo-Saharan.” Rilly and De Voogt (2012) argue that it was a relative of Nubian, and this view has gained much acceptance. Others would be happy to leave it as an unclassified language, due to the limited amount of data and the difficulty of interpreting it.

Minkin (unclassified, in Australia), though often listed among Australian language groups, is too poorly described to be classified. Inconclusive evidence points towards possible connections with Tangkic. (Bowern 2018: 324.)

Ongota (Birale) (in Ethiopia). Various scholars have offered varied hypotheses for the affiliation of Ongota, including that it is a language isolate. Blench (2018: 178) considers them all uncertain, but believes the proposed Afroasiatic connection to be the most likely. Here, it is left as of the present as unclassified.

Various different classifications of *Shabo* (Ethiopia) have been made, including considering it a language isolate. However, the language is unclassified, due to the small amount of available data (Blench 2018: 177).

A clarification of terms is also in order before we go on. Various terms intended to distinguish more inclusive from less inclusive groupings have been used in the lit-

erature, but they are confusing, since at the higher levels these terms very often refer to hypothesized but unsubstantiated “families” of languages. These postulated but unproven distant genetic relationships are sometimes referred to in the literature by the terms *stock*, *phylum*, and the compounding element ‘macro-’ (as in *Macro-Pennutian*, *Macro-Siouan*, etc.). Though these terms have been controversial, they need not complicate the count of the total number of language families in the world. The entities called ‘stock’, ‘phylum’, and ‘macro-’ would be bona fide language families if they could be established (demonstrated) and would then be counted among the world’s language families, but they are not language families if the evidence presented in their support is insufficient to show that the languages involved are related. Therefore, any entity whose component languages or language groups are definitely related, regardless of its age or of the number of branches it may have, is simply a language family. Proposed but unsubstantiated hypotheses of relationships are not language families but rather are just ‘proposed (postulated) distant genetic relationships’ and do not count as language families.

The question of how to determine whether languages not yet known to be related to one another may be distantly related is much debated. There is disagreement about the total number of independent language families in the world because of differing opinions about the status of proposed but unconfirmed possible distant genetic relationships (see Campbell and Poser 2008 for the methods for establishing kinship among distantly related languages and for evaluating hypotheses of remote relationship).

It should be pointed out, also, that there are many more language isolates than most people, including most linguists, are aware of – 162 total. That is 39.9% of the world’s c.406 language families. (See Campbell 2018 for details.) Of these language isolates, 60 are extinct (37%). Many others are highly endangered.

4. Complications and additional considerations

In addressing our ultimate question —the number of language families in the world— we also need to take into consideration several sorts of “unknowns,” to which I turn now.

4.1. Uncontacted groups

One reason that the exact number of distinct language families is not known has to do with the “uncontacted” peoples around the world, especially in Amazonia. It was estimated in 2013 that there were some 100 uncontacted groups around the world, most in South America, but some also in Papua New Guinea, Central Africa, and elsewhere (Holmes 2013). However, the number of uncontacted groups is not known, and reports and estimates vary greatly. For example, the website of FUNAI (Fundação Nacional do Índio [National Indian Foundation] in Brazil) reports that currently 107 groups of *índios isolados* (isolated indigenous groups) are registered in Amazonia (www.funai.gov.br/index.php/nossas-acoas/povos-indigenas-isolados-e-de-recente-contato), though officials at FUNAI (personal communication) indicate that they are currently tracking only 27 totally uncontacted groups. The numbers

others cite for Brazil vary from between 40 and 60 uncontacted tribes, and about 100 worldwide (see for example Holmes 2013).

The language spoken by a few of these groups is known (or strongly suspected), but in most cases it is not known whether the uncontacted groups speak varieties of already identified languages, languages currently unknown but which belong to known language families, or languages representing as yet unknown language families (possibly language isolates).

4.2. Unclassified languages

As mentioned above, it is necessary to distinguish unclassified languages from language isolates and language families. (See above for examples.) Because we are unable to classify a good number of languages that fall into this class, it is simply not possible currently to know the number of independent language families in the world.

5. Why is answering the question of how many language families there are in the world so difficult?

The answer to this question is complicated by a number of theoretical and methodological questions fundamental to historical linguistics, to which I now turn.

5.1. Is the family tree model flawed?

Some scholars have taken a skeptical view of the validity of the family tree (*Stammbaum*) (and the applicability of the comparative method) stemming from beliefs about the extent of convergence or confounding possible in language contact situations. This raises the question, can linguistic diffusion be a serious challenge to determining linguistic genealogical relationships? Or, put differently, how successful in difficult cases can we be at distinguishing inheritance from borrowing?

Jakobson (1938) offered a solution to the old debate about the possibility of multiple origins for a single language, that is, to the question about the utility of the family tree model in situations of areal diffusion: “La similitude de structure ne s’oppose donc pas, mais se superpose à la ‘parenté originaire’ des langues” (Jakobson 1949 [1938]: 353). He called for adequate description of shared traits without premature generalizations about whether they owe their explanation to a genetic relationship or diffusion. This is sound advice. It is not possible adequately to understand diffusion fully without knowing the genetic affiliation of the languages involved and vice versa, it is not possible to account fully for what is inherited without proper attention to what is diffused. That is, it is not two distinct, opposed, and antagonistic points of view that are involved, but rather both are needed; they work in concert.

Some who question the family tree model seem not to understand this. For example, we see declarations of the following sort:

The original motivation of both [areal linguistics and language typology] was the insufficiency of the genetic *Stammbaum* model for the study of relationships among languages (Dahl 2001: 1456).

Areal linguistics was originally inspired by the insufficiency of genetic relationships as an explanation for similarities between languages, in particular, by the recognition of grammatical and phonological similarities which were due to language contact (Dahl 2001: 1457).

There are misconceptions here. The goal of the historical linguist is to answer the question, what happened?, whether it be due to inheritance, diffusion, or a combination of both. To answer this question, both the inherited and the diffused must be dealt with. To test any hypothesis of genetic inheritance, it is necessary to demonstrate that it fits the facts better than alternative possible explanations, borrowing being principal among possible alternatives. Similarly, for any hypothesis of borrowing, it is necessary to demonstrate that other possible explanations do not provide a better answer, the possibility of inheritance from a common ancestor being a possible alternative explanation. Many of the errors seen today in both proposals of distant genetic relationships and in proposals that champion diffusion stem from not testing other possible explanations sufficiently before drawing conclusions in particular cases.

Some scholars in their enthusiasm for diffusion and convergence in language contact go so far as to call for alternative models and methods:

The family tree model, while appropriate and useful in many circumstances, is not applicable everywhere and cannot explain every type of relationship between languages. We need a more inclusive model, which integrates together the ideas of the family tree and of diffusion area (Dixon 1997: 28).

To reconstruct the history of a language adequately, a model is needed which is significantly more sophisticated than the family tree based on the use of the comparative method. It needs to incorporate the diffusion and layering process as well as other language-contact phenomena such as convergence, metatypy and hybridization. The desideratum is a synthesis of all the processes that affect language formation and development (Chappell 2001: 354).

Plainly, an alternative model is needed (Dixon 2002: 31).

However, most historical linguists see no need for a more inclusive integrative alternative model, since we already have methods and models that deal well with both the consequences of language contact and inheritance from a common ancestor. Inheritance and diffusion have always been of crucial importance in historical linguistics. Indeed, the comparative method is not at odds with borrowing; it is very often a major tool for detecting borrowings and thus arriving at an understanding of what is inherited and what is diffused in languages. As Watkins (2001: 59) explained, “the resilience and the power of the comparative method lies in its sensitivity to similarity due both to genetic filiation and areal diffusion alike.” For example, Hübschmann (1875) demonstrated this “when he proved that Armenian was a separate branch of Indo-European, and not a dialect of Iranian as previously thought” (Watkins 2001: 59). Armenian underwent massive influence from Iranian, but it was the application of the comparative method that revealed this as diffusion and not inheritance, and distinguished the borrowed Iranian elements from the native Armenian ones. The comparative method has been valuable in understanding the linguistic areas (diffusion zones) in which Indo-European languages participate: the Balkans, Baltic, South Asia (Indian subcontinent), and Anatolian linguistic areas. That both inherit-

ance and diffusion can be tackled with the comparative method has been shown time and again (cf. Watkins 2001).

In some cases, however, it is impossible to distinguish inheritance from diffusion. Some see this as a major problem for the comparative method and as a challenge to the family tree model (e.g. Dixon 1997, 2002). However, this view is mistaken. It is not a flaw in the comparative method if in some cases we cannot reconstruct everything due to lack of evidence. The problem is with the kind and quantity of data available, not with the model or the method.

Aikhenvald and Dixon (2001: 6-7) have claimed:

The family-tree metaphor [developed for Indo-European] has been taken over for other parts of the world in stark form, often as the sole model for relationships between [among] languages ... Rather than asking whether a form of family tree is appropriate to the language situation in some newly studied region, it has often been simply *assumed* that it is. What began as a metaphor has been ascribed reality, and has acted to constrain enquiry along narrow lines. This can lead at best to a partial and at worst to a mistaken statement of language relationships (See also Aikhenvald and Dixon 2001: 4, 6, Dixon 2002).

They seem to claim that traditional historical linguists believe that a mere diagram, used to reflect linguistic lines of descent, is the whole story and that these linguists confuse the diagram with the reality it is intended to reflect. Nevertheless, there is a reality which is not just a metaphor: languages can indeed be related to one another due to descent from a common ancestor and these relationships reflect an empirical reality. How one represents the reality (or hypotheses about the reality) graphically is not the issue. It is also not at issue that languages can borrow and undergo diffusion, changes which are also part of the traditional inquiry aimed at answering the question, what happened? As Sebeok (1950: 101) made clear, if some scholars limit their vision to only that which is inherited, too bad for them, but this is not an accurate characterization of what historical linguists do generally nor of the history of the field. The diagram, the family tree which attempts to depict inheritance and descent from an earlier common ancestor, is just one part of the larger story.

So what about cases where it is difficult or even impossible to determine whether shared traits are due to inheritance, diffusion, independent parallel development, or accident? The difficulty of distinguishing what is inherited from what is diffused is often mentioned by those who express doubts about the comparative method and genetic relationships among languages (cf. Aikhenvald 2001: 190-1, Aikhenvald and Dixon 2001: 1, Chappell 2001: 335, 353-4, Dahl 2001: 1456). All retrospective sciences are faced with the same problem: we do our best with the evidence on hand, and sometimes that evidence does not allow definitive answers. However, fortunately in linguistics our methods have proven successful over and over at distinguishing instances of borrowing from inheritance. Because the methods have been so successful, we do not abandon them just because the evidence at hand is insufficient in some specific instance, just as we do not give up on an automobile's ability to transport us just because on some occasion the gasoline ran out.

5.2. Why have typological comparisons sometimes led to erroneous hypotheses of language families?

One answer to this question is because typologically commonplace features have sometimes been mistakenly taken as sufficient evidence of relatedness among languages. However, comparison of grammatical traits that are typologically common and thus found with considerable frequency in the world's languages does nothing to enhance a proposal of genetic relationship.

This can be seen in the kinds of evidence first presented in support of the Altaic hypothesis. The core of the Altaic hypothesis holds that Turkic, Mongolian, and Tungusic are phylogenetically related. More extended versions include also Korean and Japanese, and sometimes Ainu in Altaic. The following are traits shared among these languages that are often cited in support of the Altaic hypothesis. Most of these shared features are commonplace typological traits which occur with frequency in unrelated languages of the world, and are therefore not strong evidence of relationship here. Several are also areal traits, shared by a number of other languages in surrounding regions, shared by diffusion and not reliable evidence of relatedness among languages.

1. Vowel harmony. Vowel harmony is typologically quite common, found with some frequency in languages around the world.

As proposed evidence of genetic relationship, it violates the sound–meaning isomorphism requirement, that generally accepted principle which permits only comparisons that involve both sound and meaning together. Similarities in sound alone (for example, the presence of tonal systems in compared languages) or in meaning alone (for example, grammatical gender in the languages compared) are not reliable, since they often develop independently of genetic relationship, due to diffusion, accident and typological tendencies (see Greenberg 1957, 1963). The principle is associated with Antoine Meillet; in Meillet's (1948 [1921]: 90) words:

Chinese and a language of Sudan or Dahomey such as Ewe, for example, may both use short and generally monosyllabic words, make contrastive use of tone, and base their grammar on word order and the use of auxiliary words, but it does not follow from this that Chinese and Ewe are related, since the concrete detail of their forms does not coincide; only coincidence of the material means of expression is probative.

Vowel harmony is also an areal trait, diffused across various families of central and northern Eurasia.

2. SOV word order, with modifiers before heads. This is also commonplace typologically, and subject to areal diffusion.
3. Postposition. Postpositions correlate with SOV word-order languages and this is, thus, not an independent feature. It is also common typologically.
4. Agglutinative. This is also common typologically, also a common trait in this linguistic area.
5. Comparatives based on adjectives in ablative case. This is not uncommon typologically and is also correlated with SOV word order.

6. Relatively simple phoneme inventories, absence of initial consonant clusters. This is another trait which is not uncommon typologically, and is also found in other languages of the area.
7. No verb 'to have' for possession – very common cross-linguistically.
8. No articles, no gender. This, too, is commonplace. It also violates the sound–meaning isomorphism requirement for evidence of genetic relationship.

5.3. Why is it that human genetics and other non-linguistic fields cannot help determine phylogenetic relationships among languages?

An important principle of comparative linguistics permits only linguistic information, and no non-linguistic considerations, as evidence of phylogenetic relationship among languages (Greenberg 1957, 1963). As Gabelentz (1891: 157) put it, “the only sure means for recognizing a [genetic] relationship lies in the languages themselves.” Shared cultural traits, mythology, folklore, technologies, and genes and human biological traits must be eliminated from arguments for linguistic kinship. The wisdom of this principle is seen in many failed proposals that were based on non-linguistic evidence. For example, some African classifications proposed that Ari (Omotic) belongs to either Nilo-Saharan or Sudanic “because the Ari people are Negroes,” that Moru and Madi belong to Sudanic because they are located in central Africa, or that Fula is Hamitic because the Fulani herd cattle, are Moslems, and are tall and Caucasoid (Fleming 1987: 207).

It is often assumed in recent work seeking correlations between languages and genes that somehow findings in human genetics may solve issues in the genealogical classification of languages. But the expectation of a direct association between language and genes, i.e. the assumption of parallel descent, co-evolution, is wrong (see Campbell 2015). Those who make linguistic-human genetic comparisons need to take seriously into account that while a person has only one set of genes (for life), a person can be multilingual, representing multiple languages, and that individuals (and communities) can abandon one language and adopt another, but people do not abandon their genes nor adopt new ones. Language shift (language replacement) is frequent; languages become extinct in populations which survive genetically. We cannot assume, a priori, that linguistic history and human biological history will correlate.

Non-linguistic considerations from archaeology, ethnohistory, human genetics, and so on can provide ideas about possible linguistic connections among the groups involved – that is, they can point towards hypotheses to be tested. However, evidence taken from these non-linguistic sources can never be probative, can never demonstrate linguistic kinship. Only linguistic evidence can show linguistic relationships.

5.4. What is the prognosis for discovering new family relationships among languages?

Most proposals of remote linguistic relationships have not been successful, usually either because the methods employed are inadequate or the evidence presented is insufficient to show that it could not be explained in other ways, or both. Nev-

ertheless, there have also been recent successful cases that demonstrate relatedness among languages whose relationship had not previously been demonstrated. These cases give us the optimism to believe that other cases will probably be demonstrated in the future. Some successful cases of distant genetic proposals, cases which have come to be established in the recent past to the satisfaction essentially of everyone, include:

- Austroasiatic* (Munda and Mon-Khmer)
- Harakbmut-Katukinan* (Adelaar 2000)
- Lule-Vilela* (see Viegas Barros 2001)
- Pama-Nyungan* – accepted by nearly all specialists
- Pano-Tacanan* (Panoan and Tacanan) (cf. Girard 1971)
- Paya (Pech)* as belonging to Chibchan (Holt 1986; cf. Constenla 1991)
- Tikuna-Yuri* (Carvalho 2009)
- Tlapanec-Otomanguean* (Suárez 1979, 1983, 1986; see Campbell 1997: 207-8, 211, 292, 296-7).
- Western Torres Island* and *Pama-Nyungan* (Alpher, O’Grady, and Bovern 2008).

Instances abound of languages whose genetic affiliation was previously unknown but which now has been clearly demonstrated in more recent times, for example, the several newly discovered *Dravidian* languages, *Austronesian* with the Formosan branches of the family firmly established, clarification of various languages belonging to *Sino-Tibetan*, among many others.

Therefore, although future work on language relatedness and language classification should naturally be approached with caution, applying proper methods, there is every reason to feel optimistic about possible future discoveries.

6. Conclusions

The imprecise number of approximately 406 independent language families (including 162 language isolates) is as close as we can come now to answering the question, how many language families are there in the world? It is anticipated, however, that this number will change. Some of the known families will probably be adequately demonstrated to be related to other known families through additional comparative work on languages for which we already have documentation, reducing the total number of language families. Other discoveries may increase the number of language families. The languages of at least some of the uncontacted groups will become known sufficiently well for them to be compared and perhaps for their linguistic affiliation to be determined. Probably some of the surviving unclassified languages will be described, revealing the languages involved sufficiently well for them to be classified. Possibly sufficient additional attestation may be discovered for some of the extinct unclassified languages to allow them to be identified and perhaps classified. Possibly some of the undeciphered scripts will be deciphered, revealing the languages they represent. Some of these “unknowns” will become “knowns,” and if the languages involved represent additional language families, then the overall number of language families in the world will change. However, appropriate, adequate methods for determining phylogenetic relationships among languages are crucial, and hypotheses of relatedness lacking in this regard will not be found persuasive.

References

- Adelaar, W. F. H., 2000, "Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos Indígenas de la Amazonia Occidental: Harakmbut y Katukina". In *Actas I Congreso de Lenguas Indígenas de Sudamérica*, ed. by Luis Miranda, Lima, 2: 219-236.
- & P. C. Muysken, 2004, *The Languages of the Andes*. Cambridge: Cambridge University Press.
- Aikhenvald, A. Y., 2001, "Areal diffusion, genetic inheritance, and problems of subgrouping: a North Arawak case study". In Aikhenvald & Dixon (eds.), 167-194.
- & R. M. W. Dixon, 2001, "Introduction". In Aikhenvald & Dixon (eds.), 1-26.
- & — (eds.), 2001, *Areal diffusion and genetic inheritance*. Oxford: Oxford University Press.
- Alpher, B., G. O'Grady & C. Bower, 2008, "Western Torres Strait language classification and development". In *Morphology and language history: In honour of Harold Koch*, ed. by Claire Bower, Bethwyn Evans, and Luisa Miceli, 15-30. Amsterdam: John Benjamins.
- Blench, R., 2018, "African language isolates". In Campbell (ed.), 162-192.
- Campbell, L., 1997, *American Indian languages: the historical linguistics of Native America*. Oxford: Oxford University Press.
- , 2012, "The classification of South American indigenous languages". *The Indigenous Languages of South America: A Comprehensive Guide*, ed. by Lyle Campbell and Verónica Grondona, 59-166. Berlin: Mouton de Gruyter.
- , 2015, "Do Languages and Genes Correlate?: Some Methodological Issues", *Language Dynamics and Change* 5, 202-226.
- , 2018, "Language Isolates and their history". In Campbell (ed.), 1-18.
- (ed.), 2018, *Language isolates*. Abingdon and New York: Routledge.
- & W. J. Poser, 2008, *Language classification: history and method*. Cambridge: Cambridge University Press.
- Carvalho, F. O. de, 2009, "On the genetic kinship of the languages Tikúna and Yuri", *Revista Brasileira de Linguística Antropológica* 1, 247-68.
- Catalogue of Endangered Languages*. www.endangeredlanguages.com
- Chappell, H., 2001, "Language contact and areal diffusion in Sinitic languages". In Aikhenvald & Dixon (eds.), 328-357.
- Clark, B., 2017, *A grammatical sketch of Sivia sign language*. Honolulu: University of Hawai'i Mānoa dissertation.
- Constenla Umaña, A., 1991, *Las lenguas del área intermedia. Introducción a su estudio areal*. San José: Editorial de la Universidad de Costa Rica.
- Dahl, O., 2001, "Principles of areal typology". In *Language typology and language universals: an international handbook*, ed. by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, vol. 2, 1456-70. Berlin: Mouton de Gruyter.
- Dixon, R. M. W., 1997, *The rise and fall of languages*. Cambridge: Cambridge University Press.
- , 2002, *Australian languages*. Cambridge: Cambridge University Press.
- Fleming, H. C., 1987, "Towards a definitive classification of the world's languages (Review of *A Guide to the World's Languages* by Merritt Ruhlen)", *Diachronica* 4, 59-223.

- Gabelentz, G. von der, 1891, *Die Sprachwissenschaft: ihre Aufgaben, Methoden, und bisherigen Ergebnisse*. Leipzig: T. O. Weigel Nachfolger.
- Girard, V., 1971, *Proto-Takanan phonology*. (University of California Publications in Linguistics, vol. 70.) Berkeley: University of California Press.
- Greenberg, J. H., 1957, *Essays in linguistics*. Chicago: University of Chicago Press.
- , 1963, *Languages of Africa*. (Publications of the Research Center in Anthropology, Folklore, and Linguistics no. 25.) Bloomington: Indiana University.
- Güldemann, T., 2004, “Reconstruction Through ‘de-construction’: The Marking of Person, Gender and Number in the Khoe Family and Kwadi”, *Diachronica* 21, 251-306.
- , 2008, “A Linguist’s View: Khoe-Kwadi Speakers as the Earliest Food-Producers of Southern Africa”, *Southern African Humanities* 20, 93-132.
- Holmes, B., 2013, “How many uncontacted tribes are there left in the world?”, *New Scientist*, August 22, 2013. (<https://www.newscientist.com/article/dn24090-how-many-uncontacted-tribes-are-left-in-the-world/>.)
- Holt, D., 1986, *History of the Paya sound system*. Ph.D. dissertation, University of California at Los Angeles.
- Hübschmann, H., 1875, “Über die Stellung des armenischen im Kreise der indogermanischen Sprachen”, *Zeitschrift für Vergleichende Sprachforschung* 23, 5-42.
- Jakobson, R., 1938, “Sur la théorie des affinités phonologiques entre les langues”. In *Actes du quatrième congrès international de linguistes* (tenu à Copenhague du 27 août 1^{er} septembre, 1936), 48-58. (Reprinted, 1949, as an appendix to: *Principes de phonologie*, by N. S. Troubetzkoy, 351-365. Paris: Klincksieck.)
- Meillet, A., 1948 [1921, 1914], “Le problème de la parenté des langues”. In *Linguistique historique et linguistique générale*. Paris: Champion. [Original, *Rivista di Scienza* 15(35):3.]
- Michalowski, P., 2018, “Ancient Near Eastern and European isolates”, in Campbell (ed.), 19-58.
- Mithun, M., 2018, “Language isolates of North America”. In Campbell (ed.), 260-286.
- Mufwene, S. S., 1996, “The founder principle in creole genesis”, *Diachronica* 13, 83-134.
- Rilly, C. & A. de Voogt, 2012, *The Meroitic Language and Writing System*. Cambridge: Cambridge University Press.
- Sebeok, T. A., 1950, “The importance of areal linguistics in Uralic studies”, *Memoires de la Société Finno-Ougrienne* 98, 99-106.
- Seifart, F. & H. Hammarström, 2018, “Language isolates in South America”. In Campbell (ed.), 260-286.
- Suarez, J., 1979, “Observaciones sobre la evolución fonológica del tlapaneco”, *Anales de Antropología* 16, 371-386.
- , 1983, *La lengua tlapaneca de Malinaltepec*. Mexico: UNAM.
- , 1986, “Elementos gramaticales otomangues en tlapaneco”. In *Language in global perspective: papers in honor of the fiftieth anniversary of the SIL, 1935-1985*, ed. Benjamin Elson, 267-84. Dallas: SIL.

- Thomason, S. G. & T. Kaufman, 1988, *Language contact, creolization, and genetic linguistics*. Berkeley and Los Angeles: University of California Press.
- Viegas Barros, J. P., 2001, “Evidencias de la relación genética lule-vilela”, *LIAMES* 1, 107-126.
- Watkins, C., 2001, “An Indo-European linguistic area and its characteristics: ancient Anatolia. Areal diffusion as a challenge to the comparative method?”. In Aikhenvald & Dixon (eds.), 44-63.

Lyle Campbell

Dept of Linguistics, University of Hawai'i Manoa
569 Moore Hall, 1890 East-West Road
Honolulu, HI 96822 USA
lylecamp@hawaii.edu